Consulting Assistance on Economic Reform II

DISCUSSION PAPERS

Discussion Paper 57

Infrastructure and Long Run Economic Growth[1]

David Canning

Queen=s University of Belfast,
Harvard University

and

Peter Pedroni

Indiana University

December 1999

JEL Classifications: O1, H4.

[1] The third of three technical papers under the research entitled ATelecommunications, Information Technology, and Economic Growth@ funded by the USAID CAER II project, whose assistance is gratefully acknowledged. This paper has also been funded in part by the World Bank. The views and interpretations expressed in this paper are those of the authors and should not be attributed to either the World Bank or USAID. We would like to thank Esra Bennathan, Ed Buffie, Eric Leeper, Gary Madden, and Orest Koropecky for their comments and advice. We also thank seminar participants at the North American Summer Meetings of the Econometric Society, the Midwest Macro Meetings at the University of Pittsburgh, and the Indiana University macroeconomics workshop.

Table of Contents

List of Tables

Abstract

We investigate the long run consequences of infrastructure provision on per capita income in a panel of countries over the period 1950-1992. Simple tests are devised for the existence and sign of the long run impact of infrastructure on income allowing for non-stationarity and cointegration in the time series, and heterogeneity in both the short run and long run responses across countries. We find a great deal of heterogeneity in the results across countries. Our results indicate that telephones and paved roads are provided at the growth maximizing level *on average*, but are under supplied in some countries and over supplied in others. In contrast, we find evidence that electricity generating capacity is under provided *on average*.

1. Introduction

Beginning with Aschauer (1989a, 1989b, 1989c), there has been a resurgence in the debate about the productivity effects of infrastructure. This debate is reviewed in the World Bank's *World Development Report* (1994) which finds a large range of empirical results on the importance of infrastructure for economic growth, with estimates ranging from no effect, to rates of return in excess of 100% per annum. Gramlich (1994) surveys the empirical literature and emphasizes the difficulties of accurately determining the contribution of infrastructure to growth.

Following Aschauer's early work, which found evidence for large rates of return, there have been a great number of studies using national and international data that support the result that infrastructure is important to growth. For example, using cross country data, Easterly and Rebelo (1993) find a positive effect of investment in transport and communication on economic growth. Canning, Fay, and Perotti (1994) find a positive effect of telephones on economic growth, and Sanchez-Robles (1998) finds a positive impact of road length and electricity generating capacity on subsequent economic growth. In contrast, some studies—Hulten and Schwab (1991), Tatom (1991, 1993a, 1993b), Holtz-Eakin (1994), Holtz-Eakin and Schwartz (1995), and Garcia-Mila, McGuire, and Porter (1996)—suggest that there is little evidence of an effect from infrastructure to income growth in a panel of U.S. state level data[2], particularly when fixed effects are included.

Our approach is based on the growth model of Barro (1990). Infrastructure capital is an input into aggregate production, but it comes at the cost of reduced investment in other types of capital. In this approach there is an optimal level of infrastructure which maximizes the growth rate; if infrastructure levels are set too high, they divert investment away from other capital such that income growth is reduced. We wish to investigate whether infrastructure levels are below or above their growth maximizing levels.

Using the Barro model we derive a simple "reduced form" relationship between income per capita and infrastructure stocks per capita. Below the growth maximizing infrastructure level, positive shocks to infrastructure will tend to increase the level of output, while above the optimal level, positive infrastructure shocks will tend to reduce the level of output. We use this to identify where each country's infrastructure stock stands relative to the growth maximizing level. Kocherlakota and Yi (1996, 1997) use a similar model to study the relationship between shocks to public capital and subsequent changes in gross domestic product (GDP) in the United States and the United Kingdom over the last 100 years.

Our approach to estimating the empirical relationship between infrastructure stocks and income per capita has three important features. First, we use physical measures of infrastructure, kilometers of paved roads, kilowatts of electricity generating capacity, and number of telephones. Using physical measures of infrastructure may be better than using stock estimates that have been computed by aggregating investment series. While simple physical measures do not correct for

[2] By panel of data, we mean that the data are collected over time for multiple individuals, countries or states.

quality, monetary investment in infrastructure may be a very poor guide to the amount of infrastructure capital produced. Furthermore, prices for infrastructure capital vary widely across countries, and government investment may be very inefficient, particularly in developing countries (see Pritchett (1996)).

Second, we argue that there is considerable empirical evidence that favors the presence of unit roots in per capita GDP for the types of panel data that we employ. Recent evidence of this nature includes among others Cheung and Lai (1999), Lee, Pesaran, and Smith (1997) and Pedroni (1998a, b). Such unit roots in per capita GDP, and any unit roots in the infrastructure data, are often removed by taking first differences.[3] But this may ignore evidence of a long run relationship in the data if the series are cointegrated. Indeed, our empirical analysis**C**as well as the growth model that we use to motivate the empirical work**C**supports the existence of such a cointegrating relationship between infrastructure and output. By exploiting this cointegrating relationship, we develop a simple approach to estimating the long run effect of infrastructure on GDP per capita. Furthermore, a central problem in estimating the effect of infrastructure on output is accounting for the direction of causality for both the long run and short run effects. Our technique allows us to isolate both the long run and short run consequences of infrastructure shocks in terms of the direction of causality. Notice that these shocks are by definition unexpected given the history of output growth.

Third, we estimate all our relationships separately for each country. It is common both in cross section studies and in panel data studies to assume that many parameters are common across countries. However, as noted by Pesaran and Smith (1995) and Lee, Pesaran, and Smith (1998), studies that assume incorrectly that certain parameters are common across countries can lead to bias and inconsistency when estimating worldwide average effects. Indeed, we find evidence for considerable heterogeneity among the key parameter estimates across countries, which suggests that directly pooling certain parameters across countries may be misleading in this case. Instead we estimate these relationships separately for each country, and test hypotheses regarding worldwide average effects by averaging only the relevant parameters across countries.

Cointegration between income per capita and infrastructure per capita means that there must exist an error correction mechanism, with at least one of the two variables adjusting to keep the long run equilibrium relationship intact. A natural way to explain such a relationship would be for some exogenous force, such as technical progress, to drive economic growth, and for

[3] A unit root implies that over time the data does not revert to a mean value. A random unit root process can be thought of as a generalization of a random walk process which allows for the possibility of higher order serial correlation. The terminology derives from the fact that the equation representing the time series properties of such data will have an eigenroot equal to one. By differencing a unit root process, the unit root is eliminated and the data becomes stationary.

infrastructure to respond to the level of GDP through a demand mechanism. As people become richer they may desire to have more infrastructure for consumption purposes.

But as this example illustrates, the existence of a cointegrating relationship, in itself, does not necessarily imply that causality runs in the opposite direction, from infrastructure to long run GDP per capita. The demand for infrastructure for pure consumption purposes could account for a long run cointegrating relationship between GDP per capita and infrastructure per capita. Indeed, we find strong evidence of an effect of GDP on the stock of infrastructure. However, using the approach developed in this paper, we are able to reject the hypothesis that the long run causality is unidirectional, from output to infrastructure. Rather, we also find evidence of a long run effect running from shocks in infrastructure to GDP per capita.

In fact, many economic variables tend to move together in a long run relationship with per capita income. If the long run equilibrium can be represented by a cointegrating relationship, the hypothesis that GDP per capita evolves independently, and that the variable in question simply follows GDP per capita in the long run, can be tested in a simple bivariate framework. This allows us to test if a variable's co-movement with GDP per capita merely represents a reaction to economic growth, or if there is also a direct effect on long run economic growth.

As we shall see, an advantage of this approach is that we can carry out the test without estimating the full structural model that underlies economic growth. This may be particularly useful in light of Levine and Renelt's (1992) findings on the lack of robustness of results in cross country studies of economic growth.

While we can derive a reduced form that links infrastructure stocks to GDP per capita, the nature of the long run relationship, and any short run dynamics, may vary across countries. For example, the parameters of the aggregate production function, such as the elasticity of output with respect to infrastructure, as well as the average savings rate and the proportion of investment devoted to infrastructure, may differ from country to country. The growth maximizing level of infrastructure will also vary across countries, depending on the relative importance of infrastructure in the aggregate production function of the particular country, and on the ability to finance infrastructure investment without distorting other types of investment. We therefore allow each country to have its own long run cointegrating relationship between infrastructure and output. In addition, the short run dynamics, which capture any lags before the effect of infrastructure investment reaches full potential, and the short run Keynesian multiplier effects, are all allowed to vary across countries. Furthermore, the importance of permitting such flexible cross country heterogeneity becomes even greater in reduced form systems with relatively few observable variables. This can be explained by considering that although some parameters might be modeled as homogeneous in a more fully specified model, the presence of latent unobserved variables can induce heterogeneity in the estimated coefficients on the observed variables.

In our model, we assume that infrastructure investment is financed out of savings, and that higher levels of investment in infrastructure are obtained at the cost of lowered investment in other forms of private capital. Barro (1990), and Turnovsky and Fisher (1995) investigate models in which public capital is financed out of taxation. However, it does not matter in our approach whether the infrastructure construction is the result of private sector decision on the composition of investment, or whether it is a public sector activity funded through taxes on savings, since both mechanisms reduce investment in other types of capital.

Our growth model accounts for two special cases, depending on the assumption we make about the returns to aggregate capital. In a neoclassical model, with diminishing returns to aggregate capital, infrastructure shocks and capital shocks in general have no effect on long run GDP per capita. In the endogenous growth model there are constant returns to aggregate capital. In this case, if the infrastructure level is *below* its most efficient level, then positive shocks to infrastructure stocks raise long run income per capita. By contrast, when the infrastructure level is *above* its most efficient level, the adverse impact of the reduced investment in other types of capital means that an increase in infrastructure reduces the long run level of income per capita. Finally, at the growth maximizing infrastructure level, the relationship between infrastructure and output has a turning point and the effect of an infrastructure shock is likely to be close to zero. We concentrate on the time series evidence in each country to determine which version of the model best describes the long run properties of the data, and whether infrastructure stocks appear to be supplied at levels that maximize the growth rate.

We do find evidence of long run impacts of infrastructure on GDP per capita in many countries. In this respect, our results can be interpreted as strengthening the case for endogenous growth models. However, it is important to note that for countries where we find no long run impact from infrastructure to GDP per capita, we cannot distinguish between the case in which infrastructure stocks are at the efficient level, and the case in which the neoclassical model with diminishing returns to capital holds.

For telephones and paved roads, the effect of an increase in provision on GDP per capita varies across countries, being positive in some but negative in others. *On average* telephones and paved roads are supplied at around the growth maximizing level, but some countries have too few (in particular, paved roads seem to be under-supplied in developing countries), while others have too many. It follows that the appropriate policy at the country level will depend on country specific studies. Our finding that some countries actually have too much infrastructure is consistent with Devarajan, Swaroop, and Zou (1996) and Ghali (1998), who find evidence of over provision of public capital in a number of developing countries. Our finding of zero net effect on average for paved roads is consistent with Fernald (1999) who shows that rates of return for roads in the United States are close to those for capital as a whole.

We find that long run effects of investment in electricity generating capacity are positive in a large number of countries, with negative effects being detected in only a few. By pooling these cross country effects, we find evidence that an increase in electricity generating capacity leads to higher income per capita in the long run, particularly for developed countries. This suggests that on average electricity may be under provided.

In many ways, our empirical work can also be seen as a development of Holtz-Eakin (1994) and Garcia-Mila, McGuire, and Porter (1996), who work with panel data and allow for fixed effects. We go further, allowing all other parameters to vary across the panel and exploiting the presence of cointegration in our data. Our evidence of Ano long run effect@ is, however, very different from theirs. Holding other inputs constant, they find no evidence of an effect for infrastructure. Taken at face value, this implies that infrastructure does not have any productive value. In our model we estimate the net effect on long run output of increasing infrastructure.

This is the productivity effect of the extra infrastructure, minus the effect on output of the reduction in other inputs due to the need to finance infrastructure investment. Consequently, when we find a zero long run effect, it provides support for the possibility that infrastructure is at its growth maximizing level, rather than indicating that infrastructure has no effect on growth.

In the next section we present a stylized growth model to motivate the empirical approach that is undertaken in this study. In particular, we derive our simple estimated relationship as a reduced form of a growth model. In section 3, we carry out panel-based unit root and cointegration tests to characterize the time series properties of our data that are relevant for our subsequent tests. Finally, in section 4, we discuss methods for determining the presence and direction of long effects between our variables, and present the results of these tests.

2. A Stylized Growth Model with Infrastructure Capital

In this section we describe briefly a simple stylized model of infrastructure and growth to motivate our econometric approach. In particular, we consider an economy in which infrastructure is used in the production of final output, and is financed by diverting investment from other uses, either by taxation of private savings, or by a private sector decision on the composition of investment.

The possibility of a long run impact from innovations in infrastructure to income is intimately related to the issue of whether the data are generated by a neoclassical growth model in which technical progress drives long run growth, or by an endogenous growth model in which shocks to capital accumulation can have a long run impact. In an exogenous growth model, shocks to the infrastructure stock can only have transitory effects, but in an endogenous growth model, shocks to infrastructure can lead to permanent changes in per capita income.

To make these relationships clear, consider the following very simple stylized model, which is adapted from Barro (1990). The simple model form is presented here for illustrative purposes; our estimation procedure actually allows for somewhat more general structures, as we will see. We begin by presenting the basic form of the model as it applies to a single representative country, and later discuss the implementation of the model in the context of a panel of countries. Thus, we assume for any one country that aggregate output Y, at time *t*, is

$$Y_t = A_t \; K_t^a \; G_t^b \; L_t^{1-a-b}$$

produced using infrastructure capital G, other capital K , and labor L, such that where A_t is total factor productivity at time *t*. Morrison and Schwartz (1996) provide evidence that infrastructure provision improves the productivity of private sector firms and does contribute to output. For simplicity we assume that the savings rate is constant and that both types of capital

$$G_{t+1} = t_t \; s \, Y_t$$

fully depreciate each period. Next period's infrastructure is a proportion of total savings, so that Investment in non-infrastructure capital is determined by

$$K_{t+1} = (1 - t_t)\ s Y_t$$

For our purposes it is irrelevant whether we regard the decision of how much to invest in infrastructure capital as being made by the public sector and financed out of taxes, or whether we consider it a decision made by the private sector about the allocation of investment between different sectors. If the composition of investment is set by the private sector, with competitive markets there is a presumption that it will be set at the efficient level. However, the presence of monopoly power or externalities in infrastructure provision means that the proportion of investment devoted to infrastructure may be inefficient.

Substituting the capital accumulation equations (2) and (3) into the production function (1) and dividing by L produces a difference equation for the evolution of per capita output

$$(Y/L)_{t+1} = A_{t+1} s^{a+b} (1 - t_t)^a\, t_t^b\, (Y/L)_t^{a+b} (L_t / L_{t+1})^{a+b}$$

To complete the model, we need to describe the evolution of technical progress A_t, the share of investment going to infrastructure t_t, and the size of the workforce L_t. We assume that each of these is determined by an exogenous stochastic process. We model the log of total factor

$$a_t = a_0 + s t + e_t$$

productivity a_t as
where $e_t = d\, e_{t-1} + w_t$ for some $0 \le d \le 1$, and w_t is a stationary random variable with $E[w_t] = 0$. Thus, log total factor productivity depends on a constant a_0, a trend rate of growth s, plus a random term that is stationary if $d < 1$ and non-stationary if $d = 1$.

We assume that the proportion of investment going to infrastructure is $t_t = \bar{t} + m_t$ where m_t is a zero mean stationary series. Finally we assume that the growth rate of population is given by $\log(L_{t+1} / L_t) = \bar{n} + n_{t+1}$, where n_t is a zero mean stationary series. We further assume that we can identify the workforce by the total population. Alternatively, we can easily weaken this to an assumption that the labor force participation rate is a stationary series. Under these

$$y_{t+1} = c + (a + b)\, y_t + v_{t+1}$$

assumptions, our difference equation can then be written in terms of log income per capita, y, as where $c = a_0 + s\ t + (a + b)(\log\ s - \bar{n})$ and
$v_{t+1} = e_{t+1} + a \log(1 - \bar{t} - m_t) + b \log(\bar{t} + m_t) - (a + b)\, n_{t+1}$.

Note that all the random terms in equation (6) are stationary, except possibly total factor productivity, e_{t+1}. According to equation (6) the process for y_t contains a unit root whenever $d = 1$ and $a + b < 1$, or $d < 1$ and $a + b = 1$. We require that one of these two mechanisms operates to explain the persistent, unit root, behavior in per capita income that we observe in the data, but we remain agnostic as to which one is appropriate for any particular country of our sample.

$$g_{t+1} = \bar{t} + \log s + y_t + m_t - n_{t+1}$$

Similarly, the process for infrastructure formation can be written in log per capita form as

$$g_{t+1} - \bar{t} - \log s - y_{t+1} = -\Delta y_{t+1} + m_t - n_{t+1}$$

We can rewrite this as

If y_t has a unit root, Δy_t is stationary, as are the remaining error terms in the relationship. In this case, *g* and *y* are cointegrated, since a linear combination of *g* and *y* produces a stationary variable. This will be true regardless of which assumption we use to generate the unit root in *y*. However, in the exogenous growth version, the driving force behind growth is technical progress, and long run infrastructure levels simply follow income levels. In the endogenous growth model, on the other hand, there is the possibility that shocks to infrastructure investment have permanent effects on the level of income.

Furthermore, the sign of this permanent effect may be positive or negative, depending on whether $\bar{t}$ has been set above or below the tax rate that maximizes expected growth. Note that expected growth is maximized when the average share of investment in infrastructure is set at the level t^* that maximizes the expected value of $a \log(1-\bar{t} + m_t) + b \log(\bar{t} + m_t)$. In general this depends on the distribution of the shocks. However, without shocks, setting $t^* = b/(a + b)$ maximizes the growth rate, as is shown by Barro (1990).[4]

We now summarize each of these results in the following proposition.

Proposition 1. *For the model specified by equations (1) through (8),*

(i) If $d = 1$ and $a + b < 1$, or if $d < 1$ and $a + b = 1$, then:

> *log per capita output, y_t, and log per capita infrastructure, g_t, will each be non-stationary and integrated of order one, but there will exist a cointegrating vector (possibly different for each country) such that some linear combination of g_t and y_t will*

[4] In Barro's model this is also the welfare maximizing infrastructure level. However, in the presence of shocks, increasing expected growth may also increase the volatility of the growth rate. If agents are risk averse, maximizing expected growth need not maximize expected welfare.

be stationary. Shocks to productivity have a long run positive effect on log per capita output y_t.

(ii) If $d = 1$ *and* $a + b < 1$, *then:*

shocks to per capita infrastructure, m_t, *have* ***no*** *long run effect on per capita output, y.*

(iii) If $d < 1$ *and* $a + b = 1$, *then:*
shocks to per capita infrastructure, m_t, *will have a* ***nonzero*** *long run effect on per capita output,* y_t. *For small shocks, the sign of this effect will be positive if* $\bar{t} < t^*$, *and negative if* $\bar{t} > t^*$.

The derivation of proposition 1 is provided in the appendix. In the neoclassical version of the model, shocks to infrastructure have no long run effect. In the endogenous growth version of the model, a positive shock to infrastructure increases income per capita when $\bar{t} < t^*$, and decreases income per capita when $\bar{t} > t^*$. It should be noted that all of our results correspond to small changes to infrastructure investment, since large changes could conceivably move the system across the optimal infrastructure level into a different regime.

Given these results for the reduced form structure of the model, we can estimate a bivariate relationship between income per capita and infrastructure stocks per capita, and test which version of the model best describes the long run properties of the data. The model described in this section represents a typical country of our data set. To apply the model to a panel of countries we assume that all variables and innovations terms in the model carry a double index *i, t* to represent the value of the variable in country *i* at time *t*. Furthermore, any parameters of the model are assumed to be indexed by an *i* subscript, so that we allow all of these to vary across countries. These include for example, the income share parameters of the production function, a and b, the savings rate *s*, the average share of infrastructure investment $\bar{t}$, and so forth. We also allow the parameter d, which represents the persistence of the technology shock, to vary across countries. Notice that this implies that we do not even require the exogenous growth specification or the endogenous growth specification of our model to apply uniformly; that is, the extent to which either of these specifications applies may vary from country to country.

In reality, the relationship between infrastructure and economic growth is likely to be more complex than the simple production function model we have set out. For example, the new economic geography (e.g. Krugman (1991), Holtz-Eakin and Lovely (1996)) places transport costs as a central determinant of the location and scale of economic activity, and of the pattern of trade. Harris (1995) analyzes the theoretical impact of communications infrastructure as opposed to transportation infrastructure.

However, all we really require for our empirical implementation is that the data be characterized by the properties described in the results of proposition 1. This characterization can be expected to apply to a broad class of models. Many of the assumptions that we use have been for ease of exposition and can be weakened without changing the result. Ultimately, the key assumption we maintain is that the process for log per capita output is sufficiently persistent to be modeled as a unit root process. Provided that log per capita output contains a unit root, the auxiliary result that we exploit, namely that log per capita infrastructure and log per capita output are cointegrated, is in turn likely to hold for a very broad set of models. It is essentially the consequence of the idea that the demand for infrastructure rises with income.

3. The Data

Our data are annual and were collected over the period 1950B1992. We use GDP per worker from the Penn World Tables 5.6 (see Summers and Heston (1991)). The infrastructure data are from Canning (1998), which gives physical infrastructure measured on an annual basis, in kilometers of paved road, kilowatts of electricity generating capacity, and number of telephones.

We deflate each variable by population so as to obtain per capita values, and then take logs of these per capita values. This means we have variables representing log GDP per capita, log paved roads per capita, log electricity generating capacity per capita, and log telephones per capita. If the services provided by the infrastructure stocks are considered a rival good, then these simple measures can be thought of as the average consumption of infrastructure services per capita.

We begin by investigating the time series properties of the data. In particular, we want to check that our data are consistent with the features implied by our model, which we intend to exploit in our tests for the sign and direction of the long run causal effects between infrastructure and output. Specifically, we begin by testing that g and y each have a unit root, but are cointegrated.

An important concern in testing for the order of cointegration is that, in small samples, unit root and cointegration tests lack power. We address this problem by taking the general form of the underlying growth model of the previous section to be applicable to all members of the panel. That is, the general growth model is treated as a reasonable description for each of the countries of the panel, but we do *not* assume that the same parameterization applies uniformly to all countries. In other words, we do not assume that the particular exogenous growth parameterization or the endogenous growth parameterization for d , a , or b applies uniformly to all countries, so that some countries may be characterized as exhibiting exogenous growth, while others may be characterized as exhibiting endogenous growth. Nor do we assume that any of the other parameters of the modelCthose which determine the reduced form cointegrating vector and those which determine the associated transitory dynamicsCare common among different countries of the panel. This means that we allow each country to have its own short run dynamics and its own long run cointegrating vectors.

In effect, the only condition that we pool for these first stage panel unit root and panel cointegration tests is that the order of integration of the individual variables, and the rank of any possible long run relationship among the variables, will be common across countries. This pooling condition allows us to substantially increase the power over traditional unit root and cointegration tests.

3.1 Testing for Unit Roots

There are several statistics which can be used to test for a unit root in panel data. Specifically, we wish to test for non-stationarity against the alternative that the variable is trend stationary, where we allow different intercepts and time trends for each country. We use the unit root test proposed by Im, Pesaran, and Shin (1995), which allows each panel member to have a different autoregressive parameter and different short run dynamics under the alternative hypothesis of trend stationarity. To carry out the unit root and cointegration tests, we select countries and time periods for each variable to construct a balanced panel, which entails a trade-off between the time span and number of countries in the sample.[5] For income per capita and electricity generating capacity we look at the period 1950**B**1992. However for telephones and paved roads we limit the periods to 1960**B**1990 and 1961**B**1990 respectively, in order to get a reasonable number of countries into the sample. When we come to look at the bivariate relationship the coverage of the data set is always the same as for the infrastructure variable. Before carrying out the tests, the data are purged of any common effects across countries by regressing each variable on a set of time dummies and taking residuals, as was suggested by Im, Pesaran, and Shin (1995).

The results of these unit root tests for each of our variables are shown in table 1. The test is based on the average of the augmented Dickey-Fuller (ADF) test statistics calculated independently for each member of the panel, with five lags to adjust for auto-correlation. The adjusted test statistics, (adjusted using the tables in Im, Pesaran, and Shin (1995)) are distributed as N(0,1) under the null of a unit root, and large negative values lead to the rejection of a unit root in favor of stationarity.

In no case can we reject the null hypothesis that every country has a unit root for the series in log levels. We then test for a unit root in first differences, though in this case the alternative hypothesis is stationarity without a trend, since any time trend in levels is removed by differencing. When we use first differences, the test statistic is negative and significant in each case. This indicates that we have stationarity in first differences and each of the four variables can be regarded as I(1), meaning that they become stationary only after differencing. The unit root test results for per capita GDP are also consistent with those reported in Lee, Pesaran, and Smith (1997) and Pedroni (1998a, b). In what follows we will proceed on the assumption that all

---

[5]As shown in the Monte Carlo studies reported in Pedroni (1997), nuisance parameters that are associated with the serial correlation properties of individual member country time series are eliminated asymptotically as T grows large relative to N. This suggests that we should give more weight to the time dimension when balancing the panel in order to avoid size distortion. The power of the tests, on the other hand, rises most dramatically with the N dimension, and rapidly approaches 100% against stationary but near unit root alternative hypotheses for the estimated residuals, even in relatively short panels.

log level variables are I(1) and all log differenced variables are I(0), meaning that they are stationary.

3.2 Testing for Cointegration

Now we turn to the question of possible cointegration between each infrastructure variable and GDP per capita. In the absence of cointegration we can first difference our data and then work with these transformed variables. However, in the presence of cointegration the first differences do not capture the long run relationships in the data and the cointegration relationship must be taken into account.

Given the possibility of reverse causality between the variables we use Pedroni's (1995, 1997) panel cointegration technique which is robust to causality running in both directions and allows for both heterogeneous cointegrating vectors and short run dynamics across countries. In

$$g_{it} = a_i + b_t + b_i y_{it} + e_{it}$$

particular, the cointegrating regression that we estimate is
where g_{it} is the log per capita infrastructure variable and y_{it} is log per capita income. The variable e_{it} represents a stationary error term. Note that we allow the slope of the cointegrating relationship to differ from unity and to vary across countries. This reflects the fact that in practice the relationship between infrastructure investment, infrastructure stocks, and income per capita may be more complex than set out in equation (2). Furthermore, this allows for the possibility that in practice, growth need not be balanced, so that the ratio of capital stocks to output need not be one. The common time dummies, b_t, capture any common worldwide effects that would tend to cause the individual country variables to move together over time. These may be either relatively short term business cycle effects, or longer run effects such as worldwide changes in technology that may affect the relative costs or benefits of infrastructure and thus the equilibrium relationship.

The residuals of this regression, e_{it}, are used to construct an ADF based group mean panel cointegration test from Pedroni (1997). The test is analogous to the Im, Pesaran, and Shin (1995) ADF unit root test. In table 2 we report the average over countries of the ADF t-test calculated from the residuals from regression (9) with a lag length of up to five years. Adjustment parameters to construct the test statistic are from Pedroni (1997, 1999), which allows for the fact that we are testing residuals from an estimated relationship rather than a true relationship. Large negative values imply stationarity of the residuals and lead to a rejection of no cointegration. As the results make clear, we reject the null of no cointegration in each of the three cases. Consequently, in what follows we will proceed on the assumption that each of our series is non-stationary, but that there is cointegration between each infrastructure variable and GDP per capita.

It follows from these results that the data appear to be compatible with either version of our model, the neoclassical growth model with persistent technology shocks or the endogenous growth model; both versions predict non-stationary variables and cointegration between infrastructure per capita and income per capita.

4. Long Run Effects: Empirical Implementation and Econometric Issues

Having established a long run relationship between infrastructure and income, we now turn to the issue of causality. In particular we are interested in whether innovations to infrastructure stocks have a long run effect on GDP per capita and what the sign of such an effect is. We begin this section by setting out tests for the presence and sign of such long run effects and then proceed to carry out these tests on our data.

To begin, since in each country the series *g* and *y* are individually non-stationary but together are cointegrated, we know from the Granger representation theorem (Engle and Granger (1987)) that these series can be represented in the form of a dynamic error correction model. To estimate the error correction form we employ a two step procedure. In the first step, we estimate the cointegrating relationship between log per capita infrastructure and log per capita output given in equation (9) for each country, using the Johansen (1988, 1991) maximum likelihood procedure. In the second step, we use this estimated cointegrating relationship to construct the disequilibrium term, $\hat{e}_{it} = g_{it} - \hat{a}_i - \hat{b}_t - \hat{b}_i \, y_{it}$ We then estimate the following error

$$\Delta g_{it} = c_{1i} + l_{1i} \, \hat{e}_{it-1} + \sum_{j=1}^{K} f_{11ij} \, \Delta g_{i,t-j} + \sum_{j=1}^{K} f_{12ij} \, \Delta y_{i,t-j} + e_{1it}$$

$$\Delta y_{it} = c_{2i} + l_{2i} \, \hat{e}_{it-1} + \sum_{j=1}^{K} f_{21ij} \, \Delta g_{i,t-j} + \sum_{j=1}^{K} f_{22ij} \, \Delta y_{i,t-j} + e_{2it}$$

correction model.

The variable e_{it} represents how far our variables are from the equilibrium relationship and the error correction mechanism estimates how this disequilibrium causes the variables to adjust towards equilibrium in order to keep the long run relationship intact. The Granger representation theorem implies that at least one of the adjustment coefficients l_{1i}, l_{2i} must be non-zero if a long run relationship between the variables is to hold.

By proposition 1, shocks to income have a persistent, positive component. Furthermore, the Granger representation theorem places restrictions on the singular long run response matrix of the moving average representation for the data in differences. This restricts the relationship between the long run response matrix and the speed of adjustment coefficients l_{1i}, l_{2i} in the error correction representation. We can exploit these two pieces of information to test for the existence and the sign of any long run causal effects running from innovations in log per capita infrastructure to log per capita output. We summarize our results in the following proposition. The derivation is presented in the appendix.

Proposition 2. *Given the conclusions of proposition (1),*

(i) The coefficient, l_2, *on the lagged equilibrium cointegrating relationship in the dynamic error correction equation for* Δy_t *is zero if, and only if, innovations to log per capita infrastructure have no long run effect on log per capita output.*

(ii) The ratio of the coefficients, $-l_2 / l_1$, *on the lagged equilibrium cointegrating relationship in the dynamic error correction equations for* Δy_t *and* Δg_t, *has the same sign as the long run effect of innovations to log per capita infrastructure on log per capita output.*

We can therefore test hypotheses about the long run effect of infrastructure on output by testing restrictions on the estimated coefficients in the dynamic error correction equations. According to proposition 2, a test for the significance of l_{2i} for any one country can be interpreted, conditional on our growth model, as a test of whether innovations to per capita infrastructure have a long run effect on per capita output, and a test for the sign of the ratio $-l_{2i} / l_{1i}$ can be interpreted as a test of the sign of the long run effect of innovations to per capita infrastructure on per capita output. Note that proposition 2 does not necessarily hold for cointegrated systems in general; the proof relies both on the Granger representation theorem and specific features of the growth model set out in section 2.

The advantage of our two step estimation procedure**C**first estimating the cointegrating relationship and then the error correction mechanism**C**is that all the variables in equation system (10) are stationary. Asymptotically, the fact that we use the estimated disequilibrium rather than the true disequilibrium in (10) does not affect the standard properties of our estimates, due to the well known superconsistency properties of the estimator of the cointegrating relationship.[6] It follows that we can carry out standard hypothesis tests on the coefficients estimated in the system.

By exploiting the cointegrating relationship we are able to summarize the long run effects of the innovations in the variables in terms of two parameters, l_{1i} and l_{2i}. This contrasts with those using the differenced variables in a stationary vector autoregressive (VAR) representation to estimate the impulse responses over long horizons. The tradeoff is that we only test for the existence and sign of long run effects, rather than measuring quantitatively the size of these effects. On the other hand, as is well known, the standard errors for VAR-based estimates of impulse responses over long horizons are notoriously large and unreliable, making inference difficult.[7] In essence, by exploiting the cointegrating relationships present in the data, and by

[6] Toda and Phillips (1993) study these properties in the context of more conventional dynamic Granger causality tests in cointegrated systems.

[7] See, for example, Faust and Leeper (1998) for a recent discussion of these issues. Furthermore, as Phillips (1998) demonstrates, inferences for such long horizon impulse responses are very sensitive to mis-specification of the underlying unit root and cointegration properties of the data.

summarizing the long run effects of our growth model in a small number of parameters, we avoid the problems of inference that are typically associated with summing sequences of impulse response coefficients over a long horizon.

Finally, it is worth noting that by summarizing the long run responses in the vector l, which is separate from the short run transitional dynamics associated with the coefficients of f in equations (10), we also obtain another important advantage. Specifically, the approach we use here is particularly amenable to exploiting the panel dimension of our data, since it allows us to easily pool information from the coefficients l_{1i} and l_{2i} *without* requiring that we pool the information associated with the transitional dynamics given by the coefficients f_i, or even the specific long steady state relationships reflected in the cointegrating relationship. These features are in contrast to more conventional approaches to estimating long run impulse responses in panels.[8]

4.1 Conventional Granger Causality Characterizations of the Data

We now turn to the empirical results of our tests. However, before implementing these tests for long run causal effects, we begin by asking a simpler question. We test whether the coefficients on lagged infrastructure changes, and the error correction adjustment parameter in the regression explaining income changes, are all zero. This is essentially a test of no effect from infrastructure shocks to income either in the short run or the long run. We also test for causality running in the other direction, from income to infrastructure. These tests correspond to the usual Granger causality tests in that they test whether one variable evolves entirely exogenously from another.

[8] See, for example, Holtz-Eakin, Newey, and Rosen (1988) for a discussion of the restrictions involved when using a panel VAR approach.

Column two of table 3 reports the percentage of countries that reject an F-test of the hypothesis of no causality at the 10% significance level. One interpretation of these results is that causality seems to occur in some countries, but not in others. However, considering the number of countries that are being examined, another possible interpretation is purely statistical. Specifically, if there really is no causality, we would expect to reject this hypothesis and accept causality in 10% of the countries, assuming we use the 10% significance level for our test. Rejection in a larger number of countries can be taken as evidence against the hypothesis that there is no causality in any country. Using this criterion, we have strong evidence in favor of causality running in both directions between each of our infrastructure variables and GDP, since we find rejections of no causality in a considerably more than 10% of countries.[9]

A test of the joint hypothesis of no causality in any country is given in column three of table 3. This is a likelihood ratio test of the hypothesis that all the relevant parameters are zero in every country. Under the null of no causality, the test statistic is distributed as chi-squared with degrees of freedom equal to the number of restriction imposed, which is given in parentheses beneath the statistic. Large values of this statistic lead to the rejection of the null hypothesis of no causality. Again, evidence supports two-way causality between GDP per capita and each of our infrastructure variables. The fact that non-causality is rejected in a significant number of countries supports the idea that the results for the likelihood test of non-causality in any country are not being driven by a small number of extreme estimates in a few countries.

4.2 Tests for the Presence of Long Run Effects

The conventional Granger causality results indicate two way feedback. However, the causality associated with this feedback may be only of a short run nature, so that innovations to infrastructure have an impact on GDP per capita from business cycle or multiplier effects that eventually die out and do not have a lasting effect on long run growth.

Therefore, we now turn to the issue of whether infrastructure investment affects long run economic growth. Because our variables are cointegrated, we have a simple test given by proposition 2 which asks whether the coefficient represents the adjustment of income to the disequilibrium term zero, and determines the sign of the ratio of the two disequilibrium terms. The fact that this type of test allows us to decompose effectively the short run and long run causal effects means that we can easily pool only the parameters that indicate the presence and sign of these long run effects, while allowing the short run dynamics to be unpooled. In this way, we can choose to pool the parameters for the long run effects along any one of a number of dimensions,

[9] Under the null of no causality, the percentage of countries rejecting at 10% significance level has an expected value of 10 with a standard deviation of $30N^{-1/2}$ (for N large). Using this distribution, the number of countries in which we reject no causality is significantly greater than expected even at the 1% level.

as we will see. Furthermore, we can easily compare the individual country results with the pooled results, which will allow us to distinguish average effects from sample-wide effects that hold for all countries.

We begin by considering the pattern of tests based on individual countries, and the distribution of the corresponding parameters across countries. The first test that we consider is a joint test of the hypothesis that the adjustment parameter λ_{2i} is zero in every country. We report the results of this test in column 1 of table 4. This likelihood ratio test provides strong evidence against the long run effect being uniformly zero among all countries, and easily rejects the null of no long run effect at the 1% significance level in each case.

Because it is clear that the parameters indicating the presence of a long run effect are not uniformly zero across countries, it becomes interesting to ask what the distributions of these parameters look like across the panel. First, we ask whether there is evidence that the parameters are homogeneous across countries. In table 5 we report the results for tests of homogeneity of the long run adjustment parameters across countries. Note that this test asks simply whether the long run adjustment parameters, λ_{1i} and λ_{2i} are homogeneous and treats all other parameters—including the cointegration vectors and short run adjustment parameters ϕ_i —as heterogeneous across countries. The test that we use for homogeneity is a Wald test. In practice,

$$\sum_{i=1}^{N}\frac{(\theta_i-\bar{\theta})^2}{Var(\theta_i)}$$

for a parameter θ , the test statistic is calculated as
where $\bar{\theta}$ is the weighted mean of the country-specific parameters (weighted by the inverse of their variances)[10]. Using this test, we reject decisively the homogeneity of λ_{2i} across countries.

---

[10] This is only a test that Install Equation Editor and double-click here to view equation. for all i, but it is easy to check that the test statistic is larger for any other test of the form Install Equation Editor and double-click here to view equation. . It follows that if we reject Install Equation Editor and double-click here to view equation. for all i, we reject Install Equation Editor and double-click here to view equation. for all i, for any choice of Install Equation Editor and double-click here to view equation.

Furthermore, when we test the ratio $-\lambda_{2i}/\lambda_{1i}$, which we call the sign parameter, we reject homogeneity for telephones and paved roads, though in this case only at the 10% significance level. However, it is interesting to note that we do not reject homogeneity of the sign parameter across countries in the case of electricity. As we will see, these results are important when we interpret our tests for the sign of the long run effects.

Given the heterogeneity of the parameter estimates across countries, we now examine the distribution of these estimates. The first part of table 6 gives the weighted means of each parameter estimate across countries, with weights given by the inverse of the estimated coefficient variance. The first column indicates that the average estimated value for λ_{2i} across countries is close to zero. If the parameters were homogeneous across countries, then one could interpret the mean here as an estimate of the common sample wide parameter value, in which case this could be taken as evidence for the absence of a long run effect running from infrastructure to growth. However, since we have clearly rejected parameter homogeneity for this case, we must realize that this interpretation is not valid. Rather, in this case, the average estimated value close to zero for λ_{2i} simply reflects the fact that there are many individual countries with both positive and negative values, which approximately cancel each other out.

To further support this interpretation, we note that individually many more countries are able to reject a value of zero for λ_{2i} than we would anticipate on the basis of sampling variation alone. For example, if the true parameters were zero in all countries, then we would expect to reject the null at the 10% level in 10% of the countries within our sample. In the bottom half of table 6 we report the percentage of countries that reject a zero value for λ_{2i}, and note that it far exceeds 10%. Consequently, given the parameter heterogeneity that we have found and the significance of the individual country results, this pooled result implies simply that the average effect is zero, and does not imply that there is an absence of a long run effect.

Since long run effects appear to be present in our sample, the next question is whether we can attribute a sign to the long run effect that runs from infrastructure to growth. Recall that for telephones and paved roads, we rejected homogeneity of the sign parameters. Furthermore, in table 6, column 2, results indicate that we cannot reject the possibility that the average values for these sign parameters are zero. Notice, again, that this suggests that the long run effects of increased provision of telephones and paved roads are zero *on average* across countries, but that there are significant nonzero long run effects in individual countries. Consequently, these results are consistent with the interpretation that the provision of telephones and paved roads impacts long run growth, but that the levels of provision vary depending on the growth maximizing value. Specifically, we see that these infrastructure types are under provided in some countries, and over provided in others. Nonetheless, *globally*—that is, *on average* within our sample—they are provided at approximately the growth maximizing levels.

By contrast, for electricity generating capacity we find a different type of result. We have already accepted the presence of long run effects for electricity generating capacity. Furthermore, recall that for this infrastructure type we do not reject homogeneity of the sign parameter across countries. This implies that it is possible to interpret the mean as an estimate of the single true parameter that holds approximately in each country. However, in this case we find it hard to determine the sign of the mean based on these estimates alone; while we estimate a positive effect, it is not significantly different from zero. Furthermore, as table 6 shows, the number of individual countries producing rejections of a zero value for the sign parameter is also not greater than we would anticipate based on pure sampling variation. But, as we will see shortly, by pooling the information regarding the sign parameter in a somewhat more amenable fashion, we are able to say something more definitive about the sign of these effects.

4.3 Tests for the Sign of the Long Run Effect

From a policy point of view, the key question is whether the long run growth effect of extra infrastructure is positive or negative. We now consider formal one-sided tests for the sign of the long run effect. We also consider an alternative formal test that pools the marginal significance levels for individual countries. First, however, it is interesting to note that for the simple point estimates, we report a positive value for the effect of telephones on long run growth in 54% of countries and of roads on long run growth in 52% of countries. For electricity generating capacity we find positive estimates in 60% of countries. While these results suggest a positive effect in a over half of the countries, the preponderance of positive sign coefficients over negative estimates is not overwhelming.

More formally, to evaluate the statistical strength of these findings, we can undertake one-sided tests where we test under the null hypothesis either a positive or negative sign for the long run effect against the alternative hypothesis that long run effect is of the opposite sign. Initially, we carry out such tests for each country on its own. Table 7 reports the percentage of countries rejecting the null for such one-sided tests. Again, we can ask how this compares to what we would anticipate on a pure sampling basis. For telephones, the number rejecting non-causality in favor of a positive effect and in favor of a negative effect are both significantly larger than would be predicted by pure sampling variation, and are approximately equal to one another. It follows that while telephones appear to have long run effects, the direction of the effect varies across countries. Again, we see that this implies that some countries are below their optimal level of provision of telephones relative to the growth maximizing levels, while others are above this optimal level.

For electricity generating capacity, there are a significant number of countries where shocks tend to have a positive effect. On the other hand, for paved roads there are a significant number of countries where the effect of an innovation to road stocks is to reduce long run income per capita. It appears that electricity generating capacity is under provided, while paved roads are over provided, in some countries. However, for both types of infrastructure, we can find countries where the estimated effect goes in the other direction, with electricity generating capacity appearing to be over provided and roads appearing to be under provided.

In light of these results, it is interesting to consider the consequences of pooling the results for these one-sided tests. One approach is to pool the estimates directly to construct test statistics based on the group mean, as we have done in the previous section. Another approach is to pool the marginal significance levels for the individual tests. For this, we use the Pearson r_l

statistic, also known as the Fisher statistic[11]. This is derived from the principle that for any hypothesis test over uncorrelated estimators, the test statistic

$$-2\sum_{i=1}^{N}\log(p_i)$$

has a chi-squared distribution with 2N degrees of freedom, where p_i is the marginal significance level, or Ap-value@ of the t-statistic for the parameter from country *i*. The Fisher statistic is a way of pooling independent information about a common parameter. By pooling the marginal significance values for the parameter estimate of the ratio $-l_{2i}/l_{1i}$ we are, in effect, pooling information regarding a hypothesized common value for these ratios. Again, the results from proposition 2 indicate that the sign of the long run effects for our model can be summarized by this ratio, so that pooling these does not imply that we are pooling any of the other adjustment parameters or cointegrating vectors in the error correction specification given by equation (10). Consequently, we apply the Fisher statistic to the pooled one-sided tests for the sign of the long run effect.

The results for these tests are reported in table 8. For completeness, we report results for each of our infrastructure categories. However, we should keep in mind that homogeneity of the sign parameter is only supported by the data in the case of electricity generating capacity. Consequently, only for electricity generating capacity do the results of the Fisher test have implications for a common worldwide parameter. In this case, we reject a zero long run effect in favor of a positive effect at the 5% level of significance. The results for the tests for the other types of infrastructure are really only informative if we accept that the long run sign parameter is homogeneous across countries, and so should probably be disregarded.

One potential source of heterogeneity in our results is that we are pooling countries at very different levels of income per capita. We can repeat our analysis after disaggregating countries into two groups, developed countries and less developed countries, based on their income per capita in 1960. We take $1400 a year as the dividing line, which splits our sample roughly into halves. For each of our groupings, we still find strong evidence of long run effects from infrastructure to economic growth. In table 9 we report the results of homogeneity tests on the long run sign parameter based on these groupings. We again find heterogeneity of results for telephones, but cannot reject homogeneity for electricity generating capacity. For paved roads, we find heterogeneity of the sign parameter for developed countries but cannot reject homogeneity for less developed countries.

Looking at weighted averages of the sign effects produces no results that are statistically different from zero. For the heterogeneous cases, therefore, it appears we have long run effects of different signs across countries within each group. Table 10 reports the Fisher statistic for one-sided sign tests in the cases where we accept homogeneity of the sign effects within the group. Electricity generating capacity still seems to have a positive effect on long run growth for

[11] See, for example, Maddala (1977). Lin and Shen (1996) discuss the power of the test.

developed countries, and for less developed countries the results suggest that increased provision of paved roads has a positive effect.

In summary, we have found strong evidence of long run effects running from infrastructure to income levels. For telephones and paved roads, the sign of this effect varies across countries, being zero on average, with a positive effect in some countries and a negative effect in others. There is also evidence that the effect of paved roads is positive in less developed countries. For electricity generating capacity, the data support a common positive effect in every country, though the evidence for this positive effect is stronger in developed countries.

It is also interesting to note that although a significant non-zero coefficient for the long run effect of infrastructure on GDP per capita is evidence for the endogenous growth version of our model, and a zero coefficient is consistent both with the neoclassical version and with the endogenous growth version with infrastructure stocks set at their efficient level. It follows that while the data support the endogenous growth version of the model in some countries, it is possible that different versions of the model hold in different countries.

The finding of Aoptimality@ on average must be interpreted with care. It says that countries are close to their own optimal level of infrastructure provision, which of course will differ from the optimal level in other countries. This optimum includes the cost of financing the infrastructure, and in effect evaluates the benefits in the existing economy. For example, in a poorly run economy, the benefits of infrastructure may be low, while the resource costs of supplying it may be high due to the inefficiencies of the public sector. This would give a very low optimal level of infrastructure. Therefore it is clear that from this perspective our results should be viewed as a partial equilibrium analysis, since we are evaluating infrastructure levels relative to an optimum that depends on the rest of the economy. Economic reform in the productive sector, or radical changes in the method of providing infrastructure, such as privatization, may very well change the optimal infrastructure level dramatically.

5. Conclusion

Infrastructure must be paid for. According to our model, there is a growth maximizing level of infrastructure above which the diversion of resources from other productive uses outweighs the gain from having more infrastructure. Below this level, increases in infrastructure provision increase long run income, while above this level increases in infrastructure reduces long run income. It follows that we can use the effect of shocks to infrastructure provision on long run income levels as a test of where a country=s infrastructure stock stands relative to its optimum level from a growth maximizing perspective. This is conceptually a very simple test because it does not rely on knowing the full structure of the system being examined.

Our results are interesting from the point of view of economic policy. Most studies ask, what is the effect of extra infrastructure when everything else is held constant? The question we are asking is the one more relevant to policy. That is, what is the net effect of more infrastructure taking into account that infrastructure construction diverts resources from other uses? Allowing for heterogeneity across countries is also very important for policy purposes; average results for groups of countries disguise large differences between countries. This points to the advantage of detailed country studies of the type employed by Fernald (1999) in order to find appropriate rates of return to infrastructure.

For telephones we find no evidence of a worldwide infrastructure shortage. We find that *on average* countries are near the growth maximizing levels of infrastructure provision, although individually many countries are over supplying while others are under supplying telephones. This distinction highlights the fact that, when determining where countries stand relative to their growth maximizing infrastructure levels, the study should be approached on a country by country basis. For paved roads we find similar results overall, but have some evidence of under provision in less developed countries.

For electricity generating capacity our results supports the view that there is under provision of electricity both on average and individually for the countries in our sample. However, the strong conclusion that electricity generating capacity tends to be under provided in all countries relies on our acceptance of the homogeneity of the sign parameter across countries for this variable.

In some ways our results are not surprising**.** The zero coefficients found in previous panel studies of the effect of infrastructure could be interpreted as saying that infrastructure is not important. We also find zero coefficients, but we interpret this to mean that infrastructure is important but is set close to its optimal level on average. If infrastructure were provided in competitive markets and there were no externalities present, this optimality result would be exactly what we would expect**.** However, in practice, infrastructure has often been supplied by the public sector, and we have the possibility of large externalities, perhaps leading to mis-allocation of resources**.** In this context it could be said that the finding of optimality, even if just on average, is more surprising.

References

Aschauer D. A. (1989a), AIs Public Expenditure Productive,@ *Journal of Monetary Economics*, Vol. 23, pp 177-200.

Aschauer D. A. (1989b), APublic Investment and Productivity Growth in the Group of Seven,@ *Economic Perspectives*, Vol. 13, 17-25.

Aschauer D. A. (1989c), ADoes Public Capital Crowd Out Private Capital?@ *Journal of Monetary Economics*, Vol. 24, pp 171-88.

Barro R. J. (1990), AGovernment Spending in a Simple Model of Endogenous Growth,@ *Journal of Political Economy*, Vol. 98, pp 103-125.

Canning D. (1998), AA Database of World Infrastructure Stocks, 1950-1995,@ *World Bank Economic Review*, Vol. 12, pp 529-547.

Canning D., Fay M., and Perotti R. (1994), AInfrastructure and Growth,@ in *International Differences in Growth Rates*, edited by M. Baldassarri, M. Paganaetto, and E.S. Phelps, New York, St. Martin=s Press.

Cheung Y-W. and Lai K. (1999), AInternational Evidence on Output Persistence from Postwar Data,@ forthcoming in *Economics Letters*.

Devarajan S., Swaroop V., And Zou H. F. (1996), AThe Composition of Public Expenditure and Economic Growth,@ *Journal of Monetary Economics*, Vol. 37, pp 313-344.

Easterly W. and Rebelo S. (1993), AFiscal Policy and Economic Growth: An Empirical Investigation,@ *Journal of Monetary Economics*, Vol. 32, pp 417-458.

Engle R. F. and Granger C. W. J. (1987), ACo-integration and Error Correction: Representation, Estimation, and Testing,@ *Econometrica*, Vol. 55, pp 251-276.

Faust, J. and Leeper E. (1998), AWhen Do Long-Run Identifying Restrictions Give Reliable Results?@ *Journal of Business and Economic Statistics*, Vol. 15, pp 345-354.

Fernald J. G. (1999), ARoads to Prosperity? Assessing the Link Between Public Capital and Productivity,@ *American Economic Review*, Vol. 89, pp 619-638.

Garcia-Mila T., McGuire T. J., and Porter R. H. (1996), AThe Effect of Public Capital in State Level Production Functions Reconsidered,@ *Review of Economics and Statistics*, pp 177-180.

Ghali K. H. (1998), APublic Investment and Private Capital Formation in a Vector Error-correction Model of Growth,@ *Applied Economics*, Vol. 30, pp 837-844.
Gramlich E. M. (1994), AInfrastructure Investment: A Review Essay,@ *Journal of Economic Literature*, Vol. XXXII, pp 1176-1196.

Harris R. G. (1995), ATrade and Communication Costs,@ *Canadian Journal of Economics*, Vol. XXVIII, pp S46-S75.

Holtz-Eakin D. (1994), APublic Sector Capital and the Productivity Puzzle,@ *Review of Economics and Statistics*, Vol 76, pp 12-21.

Holtz-Eakin D. and Lovely M. E. (1996), AScale Economies, Returns to Variety and the Productivity of Public Infrastructure,@ *Regional Science and Urban Economics*, Vol. 26, pp 105-123.

Holtz-Eakin D., Newey W., and Rosen H. S. (1988), AEstimating Vector Autoregressions with Panel Data,@ *Econometrica, Vol. 56*, pp 1371-95.

Holtz-Eakin D. and Schwartz A. E. (1995), AInfrastructure in a Structural Model of Economic Growth,@ *Regional Science and Urban Economics*, Vol. 25, pp 131-151.

Hulten C. R. and Schwab R. M. (1991), AIs There Too Little Public Capital? Infrastructure and Economic Growth,@ Discussion Paper, American Enterprise Institute.

Im K. S., Pesaran M. H., and Shin Y. (1995), ATesting for Unit Roots in Heterogeneous Panels,@ Working Paper 9526, Department of Applied Economics, Cambridge University.

Johansen S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, Vol. 12, pp 231-254.

Johansen S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Autoregressive Models,@ *Econometrica*, Vol. 59, pp 155-1580.

Kocherlakota N. R. and Yi K. M. (1996), AA Simple Time Series Test of Endogenous Versus Exogenous Growth Models: An Application to the United States,@ *Review of Economics and Statistics*, Vol. 78, pp 126-134.

Kocherlakota N. R. and Yi K. M. (1997), AIs there Endogenous Long-run Growth? Evidence from the United States and the United Kingdom,@ *Journal of Money Credit and Banking*, Vol. 29, pp 235-262.

Krugman P., (1991), AIncreasing Returns and Economic Geography,@ *Journal of Political Economy*, Vol. 99, pp 483-499.

Lee K., Pesaran M. H., and Smith R. (1997), AGrowth and Convergence in a Multicountry Empirical Stochastic Solow Model,@ *Journal of Applied Econometrics*, Vol. 12 pp 357-392.
Lee K., Pesaran M. H., and Smith R. (1998), AGrowth Empirics: A Panel Data ApproachCA Comment,@ *Quarterly Journal of Economics*, Vol. 113, pp 319-323.

Levine R. and Renelt D. (1992), AA Sensitivity Analysis of Cross Country Growth Regressions,@ *American Economic Review*, Vol. 82, pp 942-963.

Lin Y-L. and Shen W-H. (1996), AComparison of Powers of Tests of Fisher=s p-values,@ *Journal of Statistical Research*.

Maddala, G. S. (1977), *Econometrics*, Section 4.2, McGraw Hill, New York.

Morrison C. J. and Schwartz A. E. (1996), AState Infrastructure and Productive Performance,@ *American Economic Review*, Vol. 86, pp 1095-1111.

Pedroni P. (1995), APanel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, with an Application to the PPP Hypothesis,@ Indiana University Working Papers in Economics, No. 95-013.

Pedroni P. (1997), APanel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, with an Application to the PPP Hypothesis: New Results,@ mimeo, Indiana University.

Pedroni, P. (1998a), AOn the Role of Human Capital in Growth Models; Evidence from a Non-stationary Panel of Developing Countries,@ working paper, Indiana University.

Pedroni, P. (1998b), "Testing for Convergence to Common Steady States in Non-stationary Heterogeneous Panels," working paper, Indiana University.

Pedroni P. (1999), ACritical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors,@ *Oxford Bulletin of Economics and Statistics*, forthcoming.

Pesaran M. H. and Smith R. (1995), AEstimating Long Run Relationships from Dynamic Heterogeneous Panels,@ *Journal of Econometrics*, Vol. 68, pp 79-113.

Phillips P. C. B. (1998), AImpulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs,@ *Journal of Econometrics*, Vol. 83, pp 21-56.

Pritchett L. (1996), AMind Your Ps and Qs: The Cost of Public Investment Is Not the Value of Public Capital,@ Policy Research Working Paper, No. 1660, Policy Research Department, World Bank.

Sanchez-Robles B. (1998), AInfrastructure Investment and Growth: Some Empirical Evidence,@ *Contemporary Economic Policy*, Vol. XVI, pp 98-108.

Summers R. and Heston A. (1991), AThe Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988,@ *Quarterly Journal of Economics*, Vol. 106, pp 327-336.

Tatom J. A. (1991), APublic Capital and Private Sector Performance,@ *Federal Reserve Bank of Saint Louis Review*, Vol. 73, pp 3-15.

Tatom J. A. (1993a), APaved with Good Intentions; The Mythical National Infrastructure Crisis,@ Policy Analysis, Cato Institute.

Tatom J. A. (1993b), AThe Spurious Effect of Public Capital Formation on Private Sector Productivity,@ *Policy Studies Journal*, Vol. 21, pp 391-395.

Toda H. Y. and Philips P. C. B. (1993), AVector Autoregression and Causality,@ *Econometrica*, Vol. 61, pp 1367-1393.

Turnovsky S. J. and Fisher W. H. (1995), AThe Composition of Government Expenditure and Its Consequences for Macroeconomic Performance,@ *Journal of Economic Dynamics and Control*, Vol. 19, pp 747-786.

World Bank (1994), *World Development Report, Infrastructure for Development*, Washington D.C.

Mathematical Appendix

Proposition 1: *(i) Using equation (6) it is easy to show that y has unit root under either specification, and cointegration of y and g follows directly from equation (8). In equation (6), when $d = 1$, exogenous technology, e_t, follows a random walk, and innovations to productivity have a permanent effect on y even when $a + b < 0$. When $a + b = 1$, the endogenous process for output accumulation is no longer mean reverting, so that when exogenous technology is mean reverting, with $d < 1$, innovations to productivity have a permanent effect on y. Finally, since $a + b > 0$, positive innovations to productivity lead to positive long run effects.*

(ii) Shocks to infrastructure, m_t, only affect the steady state through their effect on y. But when $a + b < 0$, variations in y eventually dissipate because the parameter in the difference equation (6) is less than one.

(iii) In this case all shocks to output are permanent. The long run effect of an infrastructure

$$E(n) = a\log(1-\bar{t}-m) + b\log(\bar{t}+m)$$

shock to log output per capita is the same as the short run effect and is given by

$$\frac{\partial E(n)}{\partial m} = \frac{-a}{1-\bar{t}-m} + \frac{b}{\bar{t}+m}$$

Hence

$$\frac{\partial E(n)}{\partial m}\Big|_{m=0} > 0 \Leftrightarrow \bar{t} < t^*, \qquad \frac{\partial E(n)}{\partial m}\Big|_{m=0} < 0 \Leftrightarrow \bar{t} > t^*$$

Evaluating this at $m = 0$, and setting $t^* = b/(a+b)$, we have

It follows that for $\bar{t} < t^*$ small positive shocks to infrastructure raise output in both the short run and the long run while for $\bar{t} > t^*$ small positive shocks tend to reduce output. Q.E.D.

Proposition 2: *Let* $\Delta Z_t = F(L) e_t$ *be the stationary moving average representation for the differenced data* $\Delta Z_t = (\Delta g_t, \Delta y_t)'$ *in terms of the innovations* $e_t = (e_{1t}, e_{2t})'$*, so that*

$$F(1) = \begin{bmatrix} F(1)_{11} & F(1)_{12} \\ F(1)_{21} & F(1)_{22} \end{bmatrix}$$

represents the matrix of long run responses of the levels Z_t *to innovations in* e_t $F(1)_{ij}$. *represents the long run effect of j on i, and we are particularly interested in* $F(1)_{21}$*, the long run effect of infrastructure on output. According to the Granger representation theorem (Engle and Granger (1987)), if the individual series of* Z_t *are cointegrated, then the long run response matrix* $F(1)$ *will contain a singularity such that* $F(1)\, l = 0$*, where* $l = (l_1, l_2)'$ *is the vector of adjustment coefficients to the error correction term in the ECM representation given in equation (10). This implies* $F(1)_{21} l_1 + F(1)_{22} l_2 = 0$.

According to proposition 1, part (i) of section 2, we know that innovations to per capita productivity must have a positive long run effect on per capita output under either parameterization of the model, so that $F(1)_{22} > 0$*. Under cointegration an error correction mechanism exists, we cannot have both elements of* l *equal to zero. Combined with the restriction that* $F(1)_{22} > 0$*, this implies* $F(1)_{21} = 0$ *if and only if* $l_2 = 0$*, which establishes part (i) of the proposition.*

Furthermore, suppose $l_1 = 0$*. Since* $F(1)_{22} > 0$ *this implies* $l_2 = 0$ *which is a contradiction with the idea that there is an error correction mechanism. Hence* $l_1 \neq 0$ *and we can write* $F(1)_{21} = -\frac{l_2}{l_1} F(1)_{22}$*. The restriction* $F(1)_{22} > 0$ *implies that the ratio* $-\frac{l_2}{l_1}$ *has the same sign as* $F(1)_{21}$*, which establishes part (ii) of the proposition. Q.E.D.*

Table 1. Panel unit root tests.

Variable	Period	Number of countries	Average ADF	Test Statistic
log GDP per Capita	1950-1992	51	-2.164	-1.116
log EGC per Capita	1950-1992	43	-1.908	0.160
log TEL per Capita	1960-1990	67	-1.333	4.192
log PAV per Capita	1961-1990	42	-1.815	0.291
Δ log GDP per Capita	1951-1992	51	-2.465	-3.465***
Δ log EGC per Capita	1951-1992	43	-2.688	-4.863***
Δ log TEL per Capita	1961-1990	67	-2.172	-2.310**
Δ log PAV per Capita	1962-1990	42	-2.889	-5.992***

Note: The test statistics are distributed as N(0, 1) under the null hypothesis of non-stationarity. The statistics are constructed using small sample adjustment factors from Im, Pesaran, and Shin (1995). The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.
EGC represents kilowatts of electricity generating capacity.
TEL represents the number of telephones.
PAV represents kilometers of paved roads.

Table 2. Panel cointegration tests.

	Period	Countries	Average ADF	Test Statistic
Y and TEL	1960-1990	67	-2.33	-3.04***
Y and EGC	1950-1992	43	-2.28	-2.02**
Y and PAV	1961-1990	42	-2.27	-1.90**

Note: The test statistics are distributed as N(0, 1) under the null hypothesis of no cointegration.
The statistics are constructed using adjustment factors from Pedroni (1997).
The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 3. Granger causality tests.

Null Hypothesis: No Causality	Number of countries (N)	Countries rejecting null at the 10% significance level (percentage)	Full sample likelihood ratio test
Y does not cause TEL	67	37.7***	850*** (335)
Y does not cause EGC	43	51.2***	504*** (215)
Y does not cause PAV	42	45.2***	695*** (210)
TEL does not cause Y	67	46.3***	801*** (335)
EGC does not cause Y	43	30.2***	368*** (215)
PAV does not cause Y	42	42.9***	424*** (210)

Note: Under the null hypothesis of a parameter value of zero in every country, the percentage rejecting at the 10% significance level has an expected value 10 with a standard error of $30N^{-1/2}$. The likelihood ratio test is distributed as chi-squared with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 4. Tests for presence of long run effects.

Null Hypothesis: No long run effects from infrastructure to income.

	Test of l_2
	Likelihood Ratio Test
TEL to Y	325*** (67)
EGC to Y	164*** (43)
PAV to Y	211*** (42)

Note: All test statistics are distributed as chi-squared under the null hypothesis, with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 5. Tests of parameter homogeneity for long run effects across countries.

Null Hypothesis: Homogeneity of parameters across countries.

	Test of l_2 Wald Test	Test of $-l_2/l_1$ Wald Test
TEL to Y	232*** (67)	101*** (67)
EGC to Y	124*** (43)	46 (43)
PAV to Y	153*** (42)	57* (42)

Note: Test statistics are distributed as chi-squared under the null hypothesis, with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 6. Distribution of parameters.

Group Mean (weighted)	$\bar{l}_2$	$-\overline{l_2 / l_1}$
TEL to Y	0.007	-0.014
	(0.006)	(0.023)
EGC to Y	0.0014*	0.024
	(0.0007)	(0.028)
PAV to Y	0.0015	0.027
	(0.0011)	(0.061)

(Standard errors in parentheses)

Percentage of countries rejecting a zero parameter at the 10% significance level

Null Hypothesis:	$l_{2i} = 0$	$-l_{2i} / l_{1i} = 0$
TEL to Y N=67	47.6***	14.9*
EGC to Y N=43	39.5***	14.0
PAV to Y N=42	50.0***	16.7*

Note: Under the null hypothesis of a parameter value of zero in every country, the percentage rejecting at the 10% significance level has an expected value 10 with a standard error of $30N^{-1/2}$. The likelihood ratio test is distributed as chi-squared with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 7. Sign of the long run effect.

Null Hypothesis:	All effects are positive or zero (Test on $-l_2/l_1$)	All effects are negative or zero (Test on $-l_2/l_1$)
Percentage of countries rejecting the null at the 10% significance level		
TEL to Y	16.4**	16.4**
EGC to Y	9.3	16.3*
PAV to Y	21.4***	9.5

Note: Under the null hypothesis of a parameter value of zero in every country, the percentage rejecting at the 10% significance level has an expected value 10 with a standard error of $30N^{-1/2}$. The likelihood ratio test is distributed as chi-squared with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 8. Sign of the long run effect homogeneous case.

Null Hypothesis:	Effect is negative or zero (Test on $-l_2/l_1$)	Effect is negative or zero (Test on $-l_2/l_1$)
Fisher Statistic		
TEL to Y	160* (134)	148 (134)
EGC to Y	68 (68)	109** (68)
PAV to Y	94 (84)	92 (84)

Note: All test statistics are distributed as chi-squared under the null hypothesis, with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 9. Tests for parameter homogeneity among country groups.

Null Hypothesis: Homogeneity of long run sign parameter.
Test on $- l_2 / l_1$ (Wald Test)

	Less developed countries	Developed countries
TEL to Y	45* (33)	54** (33)
EGC to Y	15 (15)	32 (28)
PAV to Y	19 (17)	35* (25)

Note: All test statistics are distributed as chi-squared under the null hypothesis, with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.

Table 10. Sign of the long run effect in groups homogeneous case.

Null Hypothesis:	All Effects are positive or zero (Test on $-l_2/l_1$)	All Effects are negative or zero (Test on $-l_2/l_1$)
Fisher Statistic		
Less Developed Countries		
EGC to Y	25 (30)	35 (30)
PAV to Y	28 (34)	47* (34)
Developed Countries		
EGC to Y	43 (56)	74* (56)

Note: All test statistics are distributed as chi-squared under the null hypothesis, with the degrees of freedom given in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels.